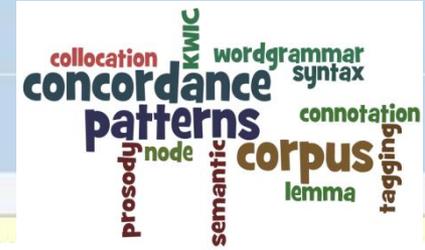


Coventry BALEAP PIM
June 21st 2014

collocation
concordance
patterns
KWIC
word
grammar
syntax
connotation
corpus
lemma
tagging
prosody
node
semantic



Using concordances in the classroom: a hands-on workshop



Learner Training

DISPLAY

WORD(S)

COLLOCATIONS

POS LIST

RANDOM

SECTIONS SHOW

FOR

SPOKEN
FICTION
MAGAZINE
NEWSPAPER
NON-ACAD

SORTING AND LIMITS

SORTING

MINIMUM

CLICK TO SEE OPTIONS

three formats: tables for relational databases, word/lemma/PoS (vertical format), or text (linear format)

Word and Phrase (analyze texts)

Word and Phrase

Word and Frequency

INTRODUCTION

[COMPARE BNC AND COCA] [COMPARE SIDE-BY-SIDE]

This website allows you to quickly and easily search the 100 million word British National Corpus (1970s-1993). The BNC was compiled by Cambridge University Press in the 1980s - early 1990s, and now exists in various versions on the web. Note that our version of the BNC is based on the CLAWS 7 tagset.

As with some other BNC interfaces, you can search for words and phrases by exact word or phrase, or by part of speech. You can also search for surrounding words (collocates) within a ten-word window (e.g. all nouns near paper, all adjectives near green).

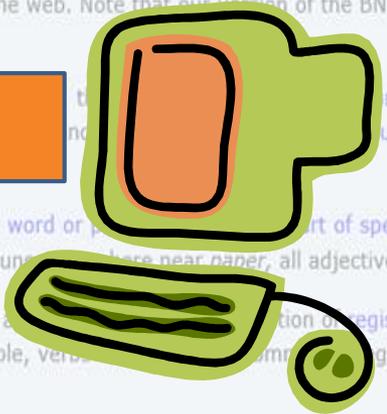
With this architecture and interface, you can also easily find the frequency of words and phrases in different registers (e.g. poetry, medical, etc). In addition, you can compare between registers -- for example, verbs that are more common in fiction than in academic writing.

As already introduced to printed concordance output in classes.

Awareness of concordancers as a self-study tool

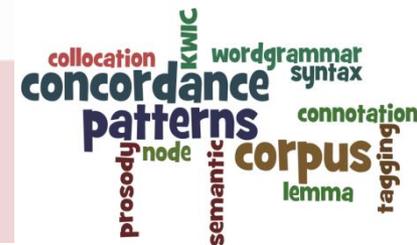
Observe - Highlight - Extract - Use Methodology

← Patterns of use: must be 2 or more



Corpus Concordance English

With precast link extractor/exporter



Student questions

CONTROLS :

Guided error correction practice

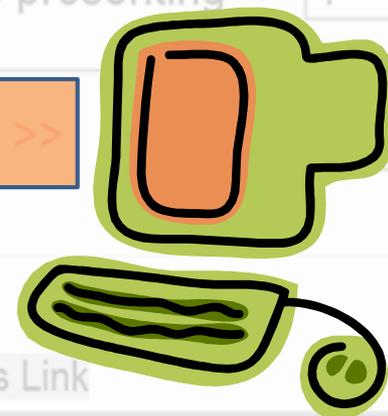
Sort By word(s) to of keyword | Line Width Number of Lines

Preparing classes

Exploring other web tools within 4 words to side.

+ Scan for any recurring word (potential colloc.) within words presenting \leq

DEMON Patterns of use: must be 2 or more >>



Link Extractor [As discussed here](#)

Encode Parameters selected as URL

Test URL

Select URL

See URL as Link

BYU-BNC: BRITISH NATIONAL CORPUS

100 MILLION WORDS, 1980s-1993

EMAIL

PASSWORD

(HELP) [LOG IN](#) ([REGISTER](#))

DISPLAY

LIST CHART KWIC COMPARE

SEARCH STRING

WORD(S)

COLLOCATES

POS LIST

[RANDOM](#)

SEARCH

RESET

SECTIONS

SHOW

1	IGNORE	2	IGNORE
----	----	----	----
SPOKEN	↑	SPOKEN	↑
FICTION	↓	FICTION	↓
MAGAZINE		MAGAZINE	
NEWSPAPER		NEWSPAPER	
NON-ACAD		NON-ACAD	

SORTING AND LIMITS

SORTING

MINIMUM 5

CLICK TO SEE OPTIONS

Full-text NEW!	Download 440 million words of full-text data for COCA (190,000 texts), or 1.8 billion words for GloWbE (1,800,000 texts). With this data, you will have the texts from the corpora on your own computer , rather than having to use the web interface. The data comes in three formats: tables for relational databases, word/lemma/PoS (vertical format), or text (linear format).
Word and Phrase (analyze texts)	Enter entire texts and see detailed frequency information on the words in the text, and create word lists based on your text. Click through the words to see detailed information on any word. Highlight phrases in your text and have it search for related phrases in COCA.
Word and Phrase (frequency lists)	Search and browse the most complete frequency dictionary of English. See detailed information (all on one page) -- definition, frequency by genre, collocates (nearby words), concordance lines, synonyms, and Wordnet-related words, all with useful links from one resource to another.
Word Frequency	You can also download lists showing the frequency of the top 60,000 lemmas by genre (and sub-genre), as well as the top 200-300 collocates (nearby words) for these lemmas (4,800,000 node/collocate pairs). There is also a free list of the top 5,000 lemmas in COCA. And now you can download the 100,000 integrated word list from COCA, COHA, BNC, and SOAP -- the largest, corrected frequency list of English.

INTRODUCTION

[Help / information / contact](#)

[\[COMPARE BNC AND COCA \]](#) [\[COMPARE SIDE-BY-SIDE \]](#)

This website allows you to quickly and easily search the 100 million word [British National Corpus](#) (1970s-1993). The BNC was originally created by [Oxford University Press](#) in the 1980s - early 1990s, and now exists in various versions on the web. Note that our version of the BNC was [recently updated](#), and it now uses the [CLAWS 7 tagset](#).

If you find this version of the BNC useful, you may also be interested in other [corpora](#) that have been created by [Mark Davies](#) of [Brigham Young University](#), including the 450 million word [Corpus of Contemporary American English](#) (1990-2010) and the 400+ million word [Corpus of Historical American English](#) (1810-2009).

As with some other BNC interfaces, you can search for words and phrases by [exact word or phrase](#), [wildcard or part of speech](#), or [combinations of these](#). You can also [search for surrounding words](#) (collocates) within a ten-word window (e.g. all nouns somewhere near *paper*, all adjectives near *woman*, or all nouns near *spin*).

With this architecture and interface, you can also easily find the frequency of words and phrases in any combination of [registers](#) that you define (spoken, academic, poetry, medical, etc). In addition, you can [compare between registers](#) -- for example, verbs that are more common in legal or medical texts, or nouns near *break* that are more common in fiction than in academic writing.

[Home](#) > [Concordancers](#) > Corpus search input (Eng)

Corpus Concordance English

With precast link extractor/exporter

[French](#) [German](#) [Spanish](#) English

Keyword(s): equals In corpus: Choose a Corpus [Corpus descriptions](#)

CONTROLS :

Sort By word(s) to of keyword | Line Width Number of Lines Gapped?

OPTION : With associated within 4 words to side.

+ Scan for any recurring word (potential colloc.) within words presenting <= times

DEMOS : <<3+4 have new sort info >> |

Link Extractor [As discussed here](#)

See URL as Link